

MODELOWANIE NEURONOWE PROCESÓW TECHNICZNYCH Z ZASTOSOWANIEM INŻYNIERII CHAOSU

Piotr Przystałka

*Instytut Podstaw Konstrukcji Maszyn, Politechnika Śląska
piotr.przystalka@polsl.pl*

Streszczenie

Artykuł dotyczy problemu modelowania neuronowego z zastosowaniem inżynierii chaosu. Główna część pracy poświęcona jest lokalnie rekurencyjnej globalnie jednokierunkowej sieci neuronowej zbudowanej z jednostek przetwarzających, dla których możliwe jest uzyskanie zachowania chaotycznego. Inżynieria chaosu wykorzystana jest w algorytmie ewolucyjnym w celu poprawy efektywności procesu uczonego. Problem wyboru wejść istotnych modelu rozwiązano, modyfikując metodę Z. Hellwiga. Izolinie kryterialne oraz wybrane metody wrażliwościowe zastosowano do poszukiwania optymalnej struktury sieci. W celu przedstawienia zalet i ograniczeń proponowanego podejścia przedstawiono wyniki modelowania neuronowego z wykorzystaniem danych zgromadzonych na obiekcie rzeczywistym.

Słowa kluczowe: rekurencyjne sieci neuronowe, algorytmy ewolucyjne, obliczenia miękkie, teoria i inżynieria chaosu.

NEURAL MODELING OF TECHNICAL PROCESSES WITH THE USE OF CHAOS ENGINEERING

Summary

The paper deals with the problem of neural modeling with the use of chaos engineering. The main part of the paper is focused on a locally recurrent neural network that is composed of complex dynamic neural units for which chaotic behaviour can be obtained. Chaos engineering is incorporated into the evolutionary algorithm in order to improve the efficiency of the tuning procedure. The problem of relevant inputs selection is solved by means of the method of extended Hellwig's coefficient of integral capacity of information. Criteria isolines and some sensitive methods are used to find the suitable architecture of a network. The merits and limitations of the proposed approach is illustrated using real-world data.

Keywords: recurrent neural networks, evolutionary algorithms, soft computing, chaos theory and chaos engineering.

1. WSTĘP

Teoria chaosu deterministycznego uznawana jest, obok teorii względności oraz mechaniki kwantowej, za jedno z trzech monumentalnych odkryć dwudziestego wieku [12,13]. Teoria chaosu w głównej mierze koncentruje swoją uwagę na analizie dynamicznych układów nieliniowych, których zachowanie jest nieregularne. Układy chaotyczne uważa się za pośrednie ogniwo

stanowiące połączenie pomiędzy układami deterministycznymi oraz stochastycznymi [18]. Dzieje się tak, ponieważ układy tego typu są deterministyczne, lecz posiadają własności charakterystyczne dla układów stochastycznych. Okazuje się, że chaos to stan dynamiki powszechnie występujący w wielu układach opisujących różnego rodzaju zjawiska (procesy).

W ostatnich latach opisano wiele rodzajów układów fizycznych, w których zaobserwowano nieregularne zachowania dynamiczne nazwane oscylacjami chaotycznymi. Przykładami tego typu układów są [16,24]: wahadło z siłą wymuszenia, płyny w pobliżu progu turbulencji, akceleratory cząstek, biologiczne modele dynamiki populacji, pobudzone komórki serca i wiele innych. Istnieje kilka głównych powodów zainteresowania tymi układami. Podstawowym czynnikiem jest wzgląd poznawczy. Układy rzeczywiste są z natury nieliniowe, dlatego wielu badaczy, próbując znaleźć przyczyny warunków generacji ruchu chaotycznego, konstruuje narzędzia umożliwiające ich analizę [1,3,28].

Z drugiej strony, ten powszechnie występujący stan dynamiki w otaczającej nas rzeczywistości jest niepożądany w niektórych sytuacjach (np. podczas pracy urządzeń elektrycznych lub elektronicznych). Pojawiło się wiele ciekawych prac w obszarze dotyczącym synchronizacji (sterowania) układów chaotycznych [19]. Dla przykładu można wymienić np. prace dotyczące stabilizacji rytmu serca [6] lub pracy lasera [22]. Innym przykładem może być zastosowanie diagramów bifurkacyjnych wyznaczanych za pomocą modelu neuronowego procesu do określenia aktualnego stanu tego procesu w celu wypracowania odpowiedniej strategii sterowania [11]. Szczegółowy przegląd zagadnień dotyczących sterowania układów chaotycznych przedstawiono np. w pracach [8,12].

W latach dziewięćdziesiątych ubiegłego wieku pojawiła się nowa koncepcja wykorzystania tego zjawiska nazwana „dobrym chaosem” (ang. good chaos) lub też „inżynierią chaosu” (ang. chaos engineering), w której uzyskanie zachowania chaotycznego w rozważanym układzie jest pożądane [1,16]. Można w tym obszarze wymienić co najmniej kilka zastosowań praktycznych, jak na przykład: wprowadzenie drgań chaotycznych o niewielkiej amplitudzie podczas procesu obróbki metali w celu uzyskania powierzchni o wysokiej jakości [17], zastosowanie systemów chaotycznych do transmisji sygnałów, przesyłu danych i kodowania informacji [25], zastosowania systemów chaotycznych jako generatorów liczb pseudolosowych [13]. Powstało również wiele ciekawych patentów urządzeń codziennego użytku (takich firm jak Sanyo, Panasonic, Goldstar), w których zastosowanie elementów teorii chaosu umożliwiło poprawę ich sprawności lub użyteczności. Przykładami takich urządzeń są pralki, zmywarki, mikrofalówki, podgrzewacze powietrza i wiele innych [8,16].

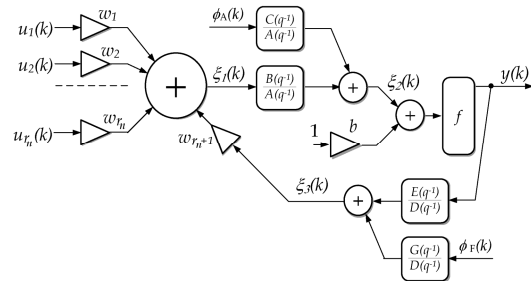
L. Zadeh w pracy [29] zaklasyfikował teorię chaosu do grupy metod stanowiących podstawę tzw. miękkich obliczeń (ang. soft computing). Jego zdaniem teoria chaosu obok podejść opartych na algorytmach ewolucyjnych, sieciach neuronowych, zbiorach rozmytych i wnioskowaniu bayesowskim może być z powodzeniem stosowana w rozwiązywaniu rzeczywistych problemów modelowania. Dzieje się tak, ponieważ metody te

bardzo dobrze nadają się do opisu zjawisk cechujących się niepewnością i nieprecyzyznością.

W niniejszym artykule przedstawiono wyniki badań autora w zakresie modelowania neuronowego z użyciem wybranych elementów teorii chaosu. Zaproponowano nowy model neuronu oraz strukturę sieci zbudowaną z takich jednostek. Opisano sposób uczenia sieci z zastosowaniem strategii dwufazowej, w której użyto układy chaotyczne do zwiększenia efektywności trenowania. Dodatkowo zaproponowano rozszerzenie metody Z. Hellwiga o miary znane z teorii chaosu do wyboru wejść relewantnych modelu neuronowego. Zaproponowany sposób modelowania zilustrowano dwoma przykładami, w których wykorzystano dane zgromadzone podczas eksploatacji rzeczywistych obiektów technicznych. Artykuł kończą wnioski przedstawiające zalety oraz ograniczenia proponowanego podejścia.

2. SIEĆ NEURONÓW DYNAMICZNYCH

Zaproponowany przez autora model neuronu pokazany jest na rys. 1 [21]. Można go uważać za uogólnienie dwóch modeli [10,20]: neuronu z filtrem o nieskończonej odpowiedzi impulsowej w bloku aktywacji oraz modelu neuronu ze sprzężeniem wyjściowym. Strukturę takiej jednostki przetwarzającej uzyskano przez wbudowanie liniowych systemów dynamicznych w blok aktywacji oraz w blok wyjściowego sprzężenia zwrotnego.



Rys. 1. Schemat ideowy przyjętego modelu neuronu

Formalny opis rozpatrywanego modelu jest następujący. Stan w złączu sumującym oblicza się z równania:

$$\xi_1(k) = \sum_{i=1}^{r_n} w_i u_i(k) + w_{(r_n+1)} \xi_3(k) \quad (1)$$

przy czym wewnętrzny stan neuronu w bloku aktywacyjnym może zostać sformułowany jako:

$$A(q^{-1})\xi_2(k) = B(q^{-1})\xi_1(k) + C(q^{-1})\phi_A(k) \quad (2)$$

oraz łączne sprzężone pobudzenie neuronu jako:

$$D(q^{-1})\xi_3(k) = E(q^{-1})y(k) + G(q^{-1})\phi_F(k) \quad (3)$$

I ostatecznie, wyjście jednostki w dyskretnej chwili k otrzymuje się z równania:

$$y(k) = \tanh\left(\frac{\xi_2(k) + b}{\alpha_f}\right) \quad (4)$$

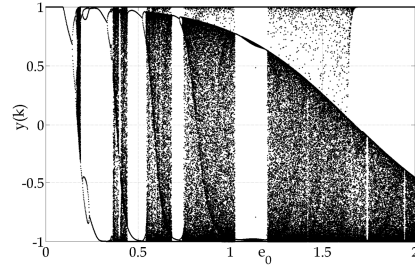
gdzie r_n określa liczbę wejść zewnętrznych neuronu, $u_i(k)$ oznacza i -te wejście neuronu w dyskretnej chwili k , w_i jest i -tą wagą neuronu, $\xi_i(k)$ oznacza stan wewnętrzny jednostki, $\phi_A(k)$, $\phi_F(k)$ reprezentują procesy losowe o założonym rozkładzie, b jest progami jednostki, α_f to parametr określający nachylenie funkcji wyjścia neuronu. Wielomiany zmiennej q^{-1} zdefiniowane są następująco:

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \dots + a_{na} q^{-na} \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \dots + b_{nb} q^{-nb} \\ C(q^{-1}) &= c_0 + c_1 q^{-1} + \dots + c_{nc} q^{-nc} \\ D(q^{-1}) &= 1 + d_1 q^{-1} + \dots + d_{nd} q^{-nd} \\ E(q^{-1}) &= e_0 + e_1 q^{-1} + \dots + e_{ne} q^{-ne} \\ G(q^{-1}) &= g_0 + g_1 q^{-1} + \dots + g_{ng} q^{-ng} \end{aligned} \quad (5)$$

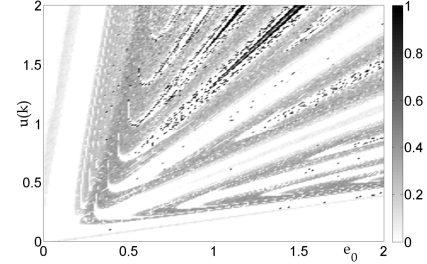
gdzie na , nb , nc , ..., ng to stopnie wielomianów zmiennej q^{-1} .

Uzyskano w ten sposób model neuronu, który posiada dwie nowe własności w stosunku do istniejących rozwiązań stosowanych obecnie w praktyce [10,20]. Po pierwsze, opracowany model neuronu jest czuły na drobne zmiany wartości jego parametrów lub parametrów otoczenia oddziałujących na niego oraz jest czuły na zmiany wartości warunków początkowych. Po drugie, wyjście neuronu dla pewnych założeń co do jego struktury może być niezdecydowane. Obie własności w sposób naturalny dziedziczy struktura lokalnie rekurencyjna. Pozwala to na dokładniejsze odwzorowanie zachowania procesów technicznych, na które oddziałują zakłócenia stacjonarne.

Ciekawą własnością dyskusowanego neuronu jest to, że przy odpowiednich założeniach (np. dla $C(q^{-1})=0$, $D(q^{-1})=1$, $G(q^{-1})=0$, $\alpha_f = 0.5$, $b_0=-1$, $w_2=5$, $a_1=-0.95$), możliwe staje się uzyskanie własności charakterystycznych dla skrajnego przypadku układów niestabilnych (tzw. układów chaotycznych). Mówi się wtedy o tak zwanym chaosie deterministycznym. Równania opisujące model neuronu są wówczas równie deterministyczne, jak w przypadku modelu neuronu wg McCullocha-Pittsa, a mimo to mogą być stosowane do opisu zachowań o charakterze losowym. Na rys. 2 i 3 rozpatrzono zachowanie neuronu z zastosowaniem diagramu bifurkacyjnego oraz analizy największego wykładnika Lapunowa. Pierwszy z wykresów uzyskano (rys. 2), zmieniając wartość parametru refrakcji ϵ_0 dla stałego wymuszenia $u(k)$. Parametr refrakcji określa wpływ poprzednich stanów wyjścia neuronu na bieżący stan wewnętrzny.



Rys. 2. Diagram bifurkacyjny dla $u(k) = 0,5$



Rys. 3. Największy wykładnik Lapunowa.

Rys. 3 przedstawia wykres konturowy obrazujący zmiany największego wykładnika Lapunowa w funkcji parametru refrakcji oraz sygnału pobudzenia. Przeprowadzona analiza potwierdza złożony charakter dynamiki dla prostego przypadku struktury neuronu uogólnionego. Uwidacznia się wiele obszarów, dla których wykładnik Lapunowa przyjmuje wartości dodatnie, co potwierdza wrażliwość zachowania neuronu ze względu na warunki początkowe.

Jednostki tego typu służą do tworzenia struktur sieci lokalnie rekurencyjnych. Przykładowa sieć neuronów dynamicznych pokazana na rys. 4 złożona jest z trzech warstw neuronów. Pierwsza warstwa pełni rolę bloku, w którym następuje wstępne przetwarzanie danych. Neurony w tej warstwie są elementami statycznymi z nieliniową funkcją wyjścia. Druga warstwa ukryta ma za zadanie odwzorowywanie pełnej dynamiki modelowanego obiektu. Neurony tej warstwy posiadają nieliniową funkcję wyjściową przy jednoczesnym stosowaniu systemów dynamicznych w blokach aktywacji i sprzężenia zwrotnego. Warstwa wyjściowa złożona jest ze statycznych jednostek z liniową funkcją wyjścia. Warstwa ta pełni rolę bloku, w którym następuje denormalizacja.

Wyjście warstwy i -tej sieci uzyskiwane jest poprzez zastosowanie liniowego lub nieliniowego operatora:

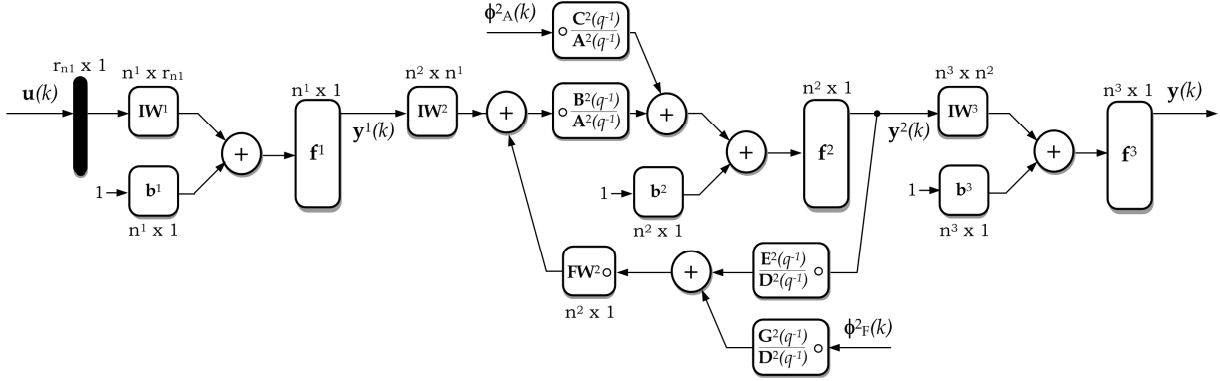
$$\mathbf{y}^i(k) = \mathbf{f}^i(\xi_2^i(k) + \mathbf{b}^i) \quad (6)$$

Wewnętrzne stany neuronów w blokach aktywacji warstwy i -tej można zapisać stosując notację wielomianową w formie wektorowej

$$\mathbf{A}^i(q^{-1}) \circ \xi_2^i(k) = \mathbf{B}^i(q^{-1}) \circ \xi_1^i(k) + \mathbf{C}^i(q^{-1}) \circ \phi_A^i(k) \quad (7)$$

i podobnie dla bloku sprzężenia zwrotnego

$$\mathbf{D}^i(q^{-1}) \circ \xi_3^i(k) = \mathbf{E}^i(q^{-1}) \circ \mathbf{y}_1^i(k) + \mathbf{G}^i(q^{-1}) \circ \phi_F^i(k) \quad (8)$$



Rys. 4. Schemat blokowy lokalnie rekurencyjnej globalnie jednokierunkowej sieci neuronowej

Pobudzenie neuronu warstwy i -tej wyznaczone jest z zastosowaniem następującego wyrażenia:

$$\xi_1^i(k) = \mathbf{IW}^i \mathbf{y}^{i-1}(k) + \mathbf{FW}^i \circ \xi_3^i(k) \quad (9)$$

gdzie \mathbf{IW}^i jest macierzą wag wejściowych neuronu, \mathbf{FW}^i jest wektorem wag sprzężenia zwrotnego, \mathbf{b}^i jest wektorem progów, $\mathbf{A}^i(q^1)$, $\mathbf{B}^i(q^1)$, ... są wektorami wielomianów zmiennej q^1 , $\xi^i(k)$ to wektor stanów i -tej warstwy, nieznane wejścia $\phi^i(k)$ reprezentowane są przez procesy losowe o założonym rozkładzie, \circ oznacza operator mnożenia tablicowego.

W pracy rozważono również problem badania stabilności uzyskiwanych modeli neuronowych. Wiadomo, że stabilność asymptotyczna sieci lokalnie rekurencyjnej jest zagwarantowana, jeżeli każdy neuron w sieci jest stabilny asymptotycznie [20]. Do określenia warunków globalnej i lokalnej stabilności asymptotycznej poszczególnych neuronów sieci użyto pośredniej metody Lapunowa. Szczegóły tych rozważań można znaleźć w monografii autora [21].

3. METODA UCZENIA

Głównym celem uczenia jest wyznaczenie wektora parametrów sieci, dla których funkcja celu przyjmuje wartość minimum. W ramach procesu uczenia zastosowano funkcję:

$$E(\omega_n) = \sum_{i=q}^N \sum_{j=1}^{n^i} |y_{ij}(\mathbf{u}_i, \omega_n) - t_{ij}|^R + \kappa \sum_{r=1}^p \frac{\omega_{nr}^2 / \omega_0^2}{1 + \omega_{nr}^2 / \omega_0^2} \quad (10)$$

gdzie ω_n jest wektorem swobodnych parametrów sieci, y_j jest j -tym wyjściem sieci, t_{ij} jest pożądanym j -tym wyjściem sieci, N to liczba wzorców uczących, p oznacza liczbę swobodnych parametrów sieci, ω_{nr} jest r -tym parametrem sieci w n -tym kroku, R , κ i ω_0 określają moc regularyzacji, n^s oznacza liczbę sygnałów wyjściowych sieci. Pierwszy składnik sumy w równaniu (10) jest

zastosowaniem następującego wyrażenia:

W celu minimalizacji funkcji (10) stosowano dwufazowe strategie optymalizacji. Najlepsze wyniki uzyskano dla strategii EA-LM będącej połączeniem algorytmu ewolucyjnego i metody Levenberga-Marquardta. Zastosowano klasyczny algorytm ewolucyjny [23] z kodowaniem rzeczywisto-liczbowym chromosomów, krzyżowaniem heurystycznym, stałą liczebnością populacji, selekcją proporcjonalną oraz sukcesją elitarną. Ponadto zaproponowano operator mutacji chaotycznej równomiernej i nierównomiernej. W pierwszym przypadku wybrany gen chromosomu zmieniany był z określonym prawdopodobieństwem mutacji r_m tak, że $\omega_j = \rho_j^i$. Chaotyczna mutacja nierównomierna realizowana była za pomocą zależności:

$$\omega^i = \omega^i + \lambda_n \rho^i \quad (11)$$

gdzie ρ^i jest wektorem o rozmiarze $p \times 1$, λ_n jest wartością zmieniającą się podczas trwania procesu ewolucyjnego zgodnie z regułą $\lambda_n = (1 - \eta n / m) \lambda_{n-1}$, przy czym $\lambda_0 = \eta$ oznacza początkową wartość wariancji rozkładu, natomiast η określa szybkość jej redukcji, m oznacza maksymalną liczbę pokoleń, n oznacza numer aktualnej epoki algorytmu. Elementy wektora ρ^i wyliczano za pomocą generatorów chaotycznych, przy czym jedne z najlepszych wyników optymalizacji uzyskiwano dla układu Hénona [3]:

$$\rho^i = \left[\begin{array}{l} \tilde{\rho}_j^i = \rho_{j-1}^i + 1 - 1.4(\tilde{\rho}_{j-1}^i)^2 \\ \rho_j^i = 0.3\tilde{\rho}_{j-1}^i \end{array} \right]_{j=1}^p \quad (12)$$

Następnie metoda optymalizacji lokalnej, startując z punktu znalezionej za pomocą poprzedniego algorytmu, doprowadza do poszukiwanego rozwiązania optymalnego zgodnie ze wzorem rekurencyjnym

$$\omega_{n+1} = \omega_n - \left[\mathbf{J}^T(\omega_n) \times \mathbf{J}(\omega_n) + \eta_n \text{diag}(\mathbf{J}^T(\omega_n) \times \mathbf{J}(\omega_n)) \right]^{-1} \times \mathbf{J}^T(\omega_n) \times \mathbf{E}(\omega_n) \quad (13)$$

gdzie jest η_n parametrem skalarnym określającym moc składnika regularyzacyjnego wyliczanego w sposób klasyczny [21], \mathbf{J} jest jacobianem funkcji celu (10), który uzyskuje się stosując metody numeryczne lub metody ze statystycznym estymatorem. Na ogół przyjmuje się niewielką liczbę epok dla algorytmu ewolucyjnego, co sprzyja odnalezieniu rozwiązania dość odległego od optymalnego, lecz na tyle bliskiego, aby metoda ucząca nie utknęła w optimum lokalnym.

4. DOBÓR WEJŚĆ I OPTYMALIZACJA STRUKTURY SIECI

W celu doboru istotnych wejść modelu autor zaproponował udoskonalenie koncepcji metody wskaźników pojemności informacyjnej (MWI), której pionierem był Z. Hellwig. W metodzie tej pojemność integralna stanowi kryterium wyboru odpowiedniego zestawu zmiennych wejściowych:

$$H_k^c = \sum_{j=1}^{L_c} \frac{v_{jk}^2}{1 + \sum_{i=1, i \neq j}^{L_c} |v_{ij}|} \quad (14)$$

gdzie v_{jk} - współczynnik określający zależność pomiędzy k -tą zmienną wyjściową, a j -tą zmienną wejściową modelu, v_{ij} współczynnik określający zależność pomiędzy i -tą i j -tą potencjalną zmienną wejściową modelu, $c = 1, 2, \dots, 2^L - 1$ oznacza numer kombinacji, L jest liczbą zmiennych, L_c jest liczbą zmiennych w c -tej kombinacji.

Wybiera się tę kombinację zmiennych wejściowych, dla której pojemność integralna jest maksymalna. W odróżnieniu od standardowej metody Z. Hellwiga jako miary zależności pomiędzy zmiennymi procesowymi stosowano tu: współczynnik korelacji Pearsona, Spearmana, Kendalla oraz informację wzajemną. Wymienione wcześniej współczynniki pozwalają na badanie zależności statycznych. Jako rozwiązanie tego problemu autor zaproponował rozszerzenie współczynników v_{ij} o miary bazujące na analizie diagramów rekurencyjnych RQA (ang. Recurrence Quantification Analysis) [14]. Aby można było wyliczyć miary RQA dla wybranych zmiennych procesowych, niezbędne jest odpowiednie ich przekształcenie. Przekształcenie rozpatrywanych zmiennych realizuje się, stosując metodę współrzędnych z opóźnieniem [30], otrzymując dwie trajektorie takie, że:

$$\begin{aligned} \mathbf{z}_{ij} &= [z_{ij} \quad z_{(i+\tau_j)j} \quad z_{(i+2\tau_j)j} \quad \dots \quad z_{(i+d\tau_j-\tau_j)j}]^T \\ \mathbf{z}_{ik} &= [z_{ik} \quad z_{(i+\tau_k)k} \quad z_{(i+2\tau_k)k} \quad \dots \quad z_{(i+d\tau_k-\tau_k)k}]^T \end{aligned} \quad (15)$$

gdzie $i = 0, 1, \dots, N - (d-1)\tau_j$ a czasy charakterystyczne τ_j i τ_k oraz wymiar zanurzenia d dobiera się za pomocą znanych metod [15]. Poprzeczny diagram rekurencyjny wyznaczany jest z zastosowaniem cechy funkcyjnej (będącej miarą rekurencji) zdefiniowanej jako:

$$[\mathbf{CR}]_{nm} = H(\varepsilon - \|\mathbf{z}_{nj} - \mathbf{z}_{mk}\|) \quad (16)$$

gdzie $i = 0, 1, \dots, N_1 = N - (d-1)\tau_j$, H to funkcja Heaviside'a, $\|\bullet\|$ jest wybraną metryką (w pracy rozważano metrykę euklidesową), próg $\varepsilon = 5\sigma$, przy czym σ określa odchylenie standardowe szumu pomiarowego występującego w rozpatrywanych danych [27]. Dla tak wyliczonej cechy funkcyjnej możliwe jest wygenerowanie diagramu rekurencyjnego oraz miar RQA. W pracy ograniczono się do miar RQA takich jak [14]:

- Wskaźnik rekurencji - odsetek punktów rekurencyjnych w diagramie:

$$rr = \frac{1}{N_1^2} \sum_{n,m=0}^{N_1} [\mathbf{CR}]_{nm} \quad (17)$$

- Determinizm - zdefiniowany jako odsetek punktów rekurencyjnych, które należą do linii diagonalnych diagramu (o długościach co najmniej l_{\min}):

$$\det = \frac{\sum_{l=l_{\min}}^{N_1} lP(l)}{\sum_{l=1}^{N_1} lP(l)} \quad (18)$$

- Entropia - entropia informacyjna Shannona wyznaczana dla rozkładu prawdopodobieństwa linii diagonalnych:

$$\text{ent} = - \sum_{l=l_{\min}}^{N_1} p(l) \ln p(l) = - \sum_{l=l_{\min}}^{N_1} \frac{P(l)}{N_1} \ln \frac{P(l)}{N_1} \quad (19)$$

- Laminarność - odsetek punktów należących do linii pionowych, których długość wynosi co najmniej l_{\min}

$$\text{lam} = \frac{\sum_{l=l_{\min}}^{N_1} lP_v(l)}{\sum_{l=1}^{N_1} lP_v(l)} \quad (20)$$

- Czas pułapkowania - jest średnią długością linii pionowych diagramu:

$$\text{tt} = \frac{\sum_{l=l_{\min}}^{N_1} lP_v(l)}{\sum_{l=l_{\min}}^{N_1} P_v(l)} \quad (21)$$

gdzie $P(l)$ jest prawdopodobieństwem znalezienia linii diagonalnej o długości l na diagramie, N_l jest całkowitą liczbą linii diagonalnych diagramu.

Dobór architektury sieci neuronowej jest jednym z czynników wpływających na zdolność uogólniania (generalizacji) tej sieci. Proponowana metoda selekcji struktury sieci jest połączeniem podejścia bazującego na izoliniach kryterialnych [23] oraz metodzie wrażliwościowej OBD (ang. Optimal Brain Damage) [4]. Pierwsza z wymienionych metod stosowana jest do wstępnego doboru struktury sieci, przy czym możliwe jest porównywanie topologii sieci o różnym typie, jak również modeli różnych klas (np. klasycznych modeli parametrycznych lub modeli neuronowo-rozmytych). Jest to podejście, w którym wykreślana jest mapa izolini dla określonych kryteriów informacyjnych (znanych z identyfikacji systemów). Izolinie reprezentują obszary, dla których wartości rozpatrywanego kryterium są stałe w funkcji złożoności modelu p oraz wartości oceny dokładności modelu Q_G . W badaniach autor stosował różne kryteria informacyjne [21], z których najważniejsze to:

- kryterium Bayesa:

$$\text{BIC} = N_G \ln(Q_G) + p \ln(N_G) \quad (22)$$

- kryterium Hannana i Quinna:

$$\text{HQC} = N_G \ln\left(\frac{Q_G}{N_G}\right) + 2p \ln(\ln(N_G)) \quad (23)$$

gdzie Q_G jest średnim względnym błędem procentowym wyznaczonym dla wzorców testowych z godnie ze wzorem

$$\text{mMAPE} = \frac{100\%}{n^s N_G} \sum_{j=1}^{n^s} \sum_{i=0}^{N_G} \left| \frac{y_{ij} - \hat{y}_{ij}}{\max(Y_j) - \min(Y_j)} \right| \quad (24)$$

przy czym N_G to liczba wzorców testowych. Na tak wykreślanych mapach nanosi się wyniki testów dla wygenerowanych modeli. Zastosowanie wykresów izolini kryterialnych oraz zasady oszczędności pozwala na wskazanie struktury najbardziej obiecującej. W kolejnym kroku dla tak wybranego modelu stosowana jest procedura OBD, której kluczowym elementem jest wyznaczenie miary istotności s_k dla każdego parametru swobodnego sieci zgodnie ze wzorem:

$$s_k = \frac{1}{2} \frac{\partial^2 E(\mathbf{w})}{\partial \omega_k^2} \omega_k^2 = \frac{1}{2} h_{kk} \omega_k^2 = \frac{1}{2} [\mathbf{J}(\mathbf{w}) \times \mathbf{J}(\mathbf{w})^T]_{kk} \omega_k^2 \quad (25)$$

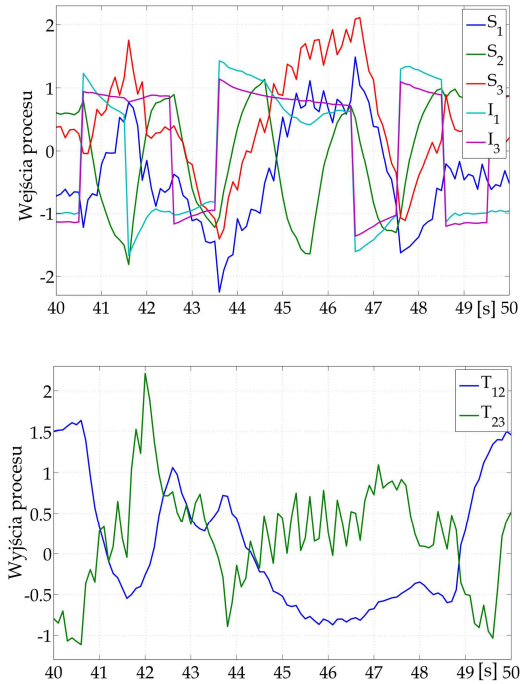
gdzie h_{kk} jest k -tym elementem diagonalnym macierzy drugich pochodnych funkcji (10). Elementy te wyznaczone są bezpośrednio lub z użyciem aproksymacji za pomocą jacobianu \mathbf{J} . Parametry, dla których uzyskano najmniejsze wartości wskaźnika s_k zostają usunięte, a model douczany jest z zastosowaniem algorytmu lokalnego wg reguły (13).

5. PRZYKŁADY ZASTOSOWANIA

5.1 MODELOWANIE PROCESU ZWIJANIA TKANINY

W celu wykazania praktycznej użyteczności zaproponowanego podejścia, przeprowadzono testy z użyciem danych, które pozyskano na pilotażowym obiekcie realizującym proces zwijania tkaniny [5]. Stanowią one część obszernego zbioru danych zgromadzonych dla celów weryfikacji i badania różnych metod odkrywania wiedzy.

Obiekt identyfikacji złożony jest z trzech rolek (rozwijającej, pociągowej i nawijającej), które napędzane są przez odpowiednio sprzęgnięte silniki prądu stałego z przekładnikami redukującymi. Wartościami mierzonymi są: prędkości obrotowe rolek – S1, S2, S3 [rad/s], naciąg pomiędzy rolkami 1 i 2 oraz 2 i 3 - T12, T23 [N], wartości zadane prądów silników 1 i 3 - I1 i I3 [A]. Każdy z silników sterowany jest w pętli sprzężenia zwrotnego z zastosowaniem regulatorów PI. Wartości zadane (prądu i prędkości obrotowej) generowane są przez odpowiedni sterownik PLC w celu utrzymania zadanej prędkości i naprężenia materiału. Na rys. 5 pokazano przebiegi rejestrowanych zmiennych procesowych. Odstęp pomiędzy kolejnymi próbkami wynosi 0.1 [s]. Dane obejmują około cztery minuty realizacji procesu (2500 próbek). Dostępne sygnały wejściowe i wyjściowe podzielono na zbiory danych trenujących oraz testowych w proporcjach odpowiednio 60% i 40%.



Rys. 5. Znormalizowane przebiegi zmiennych procesowych

W pierwszej fazie eksperymentu identyfikacyjnego utworzono dwa wzorcowe modele procesu (\diamond): model parametryczny ARX oraz model neuronowy typu

NNARX [21]. Strukturę pierwszego modelu dobierano metodą przeszukiwania systematycznego z zastosowaniem kryterium AIC oraz zasady oszczędności. Otrzymano model ARX (rys. 6 - model nr 226) o strukturze $\mathbf{NA} = [3 \ 3; 3 \ 3]$, $\mathbf{NB} = [3 \ 3 \ 3 \ 3 \ 3; 3 \ 3 \ 3 \ 3 \ 3]$ oraz $\mathbf{NK} = [0 \ 0 \ 0 \ 0 \ 0; 0 \ 0 \ 0 \ 0 \ 0]$, którego parametry dostrojono z zastosowaniem metody najmniejszych kwadratów. Dla tego modelu uzyskano wartość błędu testowania mMAPE 4%.

Strukturę drugiego modelu dobierano podobnie jak poprzednio. Najlepsze wyniki otrzymano dla modelu o strukturze $5 \rightarrow \begin{matrix} (0,1) \\ (1,2) \end{matrix} 6 \rightarrow 2$ z funkcjami tangensoidalnymi neuronów warstwy ukrytej i liniowymi warstwy wyjściowej. Przyjęto oznaczenia struktur modeli zgodnie z notacją oprogramowania MATLAB co zostało szczegółowo wyjaśnione w pracy [21]. Parametry tego modelu dostrajano algorytmem Levenberga-Marquardta z regularyzacją Bayesa, uzyskując wartość błędu testowania mMAPE $\approx 3.7\%$.

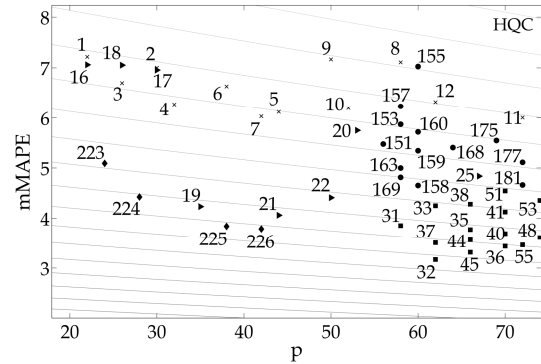
Następnym krokiem było utworzenie zbioru modeli lokalnie rekurencyjnych (,) opracowanych wg metody opisanej w rozdziale 2, modeli wielokontekstowych Jordana (x) i Elmana () [21]. Przyjęto ograniczenie, że w warstwach ukrytych neurony posiadają funkcje wyjścia w postaci $\tanh(x)$, a w warstwie wyjściowej funkcje liniową. Podczas uczenia zastosowano algorytm hybrydowy według schematu EA-LM, przy czym przyjęto następujące wartości parametrów funkcji celu (10) dla algorytmu globalnego: $R = 4$, $\mathbf{K} = 0.5$, $\boldsymbol{\omega}_0 = 1.5$ oraz dla algorytmu lokalnego: $R = 2$, $\mathbf{K} = 0$.

W pierwszym etapie treningu stosowano algorytm ewolucyjny (10 epok) z kodowaniem rzeczywistoliczbowym chromosomów, z krzyżowaniem heurystycznym (prawdopodobieństwo krzyżowania $p_k = 0.8$) z mutacją równomierną z odwzorowaniem Hénona (prawdopodobieństwo mutacji $r_m = 0.1$), stałą liczebnością populacji równą 50 osobników, gdzie populacja bazowa uzyskiwana była metodą zaproponowaną w [20], selekcją proporcjonalną oraz sukcesją elitarną. W kolejnym etapie procesu dostrajania modeli stosowano algorytm Levenberga-Marquardta (15 iteracji) z numerycznie wyznaczanym jacobianem. Uzyskane wyniki zaprezentowano w tab. 1 i na rys. 6.

Tab. 1. Oceny poprawności działania wybranych modeli neuronowych procesu

$(\mathbf{T}_{12}; \mathbf{T}_{23}) = f(\mathbf{S}_1; \mathbf{S}_2; \mathbf{S}_3; \mathbf{I}_1; \mathbf{I}_3)$					
Nr	Ozn.	Struktura	p	mMAPE	HQC
32		$5 \rightarrow 4 \begin{matrix} (2,2,0) \\ (0,2,0) \end{matrix} \rightarrow 2$	62	3.168	-5514
45		$5 \rightarrow 4 \begin{matrix} (3,2,1) \\ (0,1,0) \end{matrix} \rightarrow 2$	66	3.313	-5454
21		$5 \rightarrow 3 \begin{matrix} (1,2) \\ (1,2) \end{matrix} \rightarrow 2$	44	4.060	-5336
4	x	$5 \rightarrow 3 \begin{matrix} (1) \\ (1) \end{matrix} \rightarrow 2$	32	6.272	-4947
...

Na podstawie analizy rys. 6 możliwe było wskazanie modelu neuronowego najbardziej obiecującego. Dla tego przypadku wybrano model o numerze 32. Kolejnym etapem modelowania procesu była optymalizacja poszczególnych połączeń modelu z zastosowaniem metody OBD. Po usunięciu (wyzerowaniu) parametrów nieistotnych przeprowadzono proces douczania modelu wyłącznie z użyciem algorytmu Levenberga-Marquardta (10 iteracji).

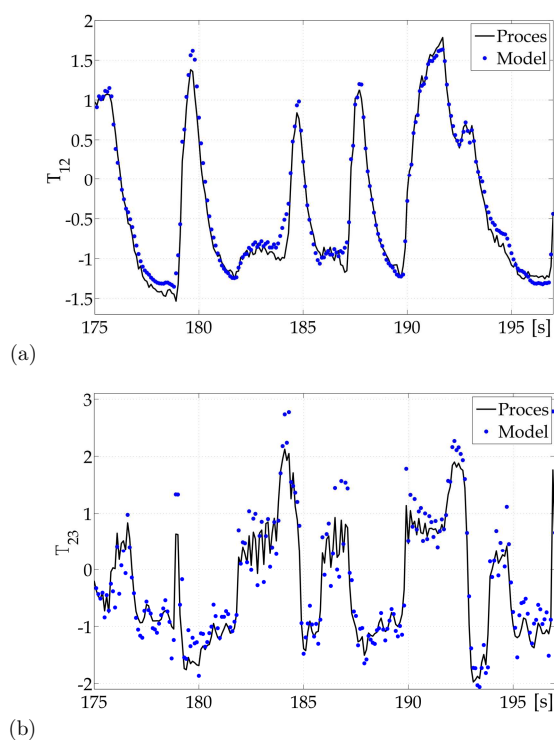


Rys. 6. Linie izokryterialne: wyniki testowania osiągnięte z użyciem wybranych modeli

Wyniki testowania modelu przed i po etapie przycinania sieci zawarto w tab. 2. Wyznaczone wartości względnego błędu predykcji RPE [21] (w odniesieniu do modeli wzorcowych ARX i NNARX) umożliwiającą obiektywną ocenę jakości odwzorowania dynamiki obiektu.

Tab. 2. Oceny poprawności działania modelu neuronowego - przed i po procesie przycinania jego struktury

Model	p	mMAPE	RPE(ARX)	RPE(NNARX)
przed przycinaniem				
32	62	3.169	0.796	0.848
...
po przycinaniu				
32 ₅₈	58	3.586	0.901	0.959
32 ₅₇	57	4.489	1.128	1.201

Rys. 7. Wyjścia z procesu i modelu neuronowego 32_s

Z rys. 7(a) i 7(b) oraz z tab. 2 wynika, że zadowalające odwzorowywanie dynamiki obiektu możliwe było z użyciem struktury lokalnie rekurencyjnej o dużo mniejszej złożoności niż w przypadku sieci typu NNARX oraz znacznie większej dokładności niż w przypadku modelu ARX.

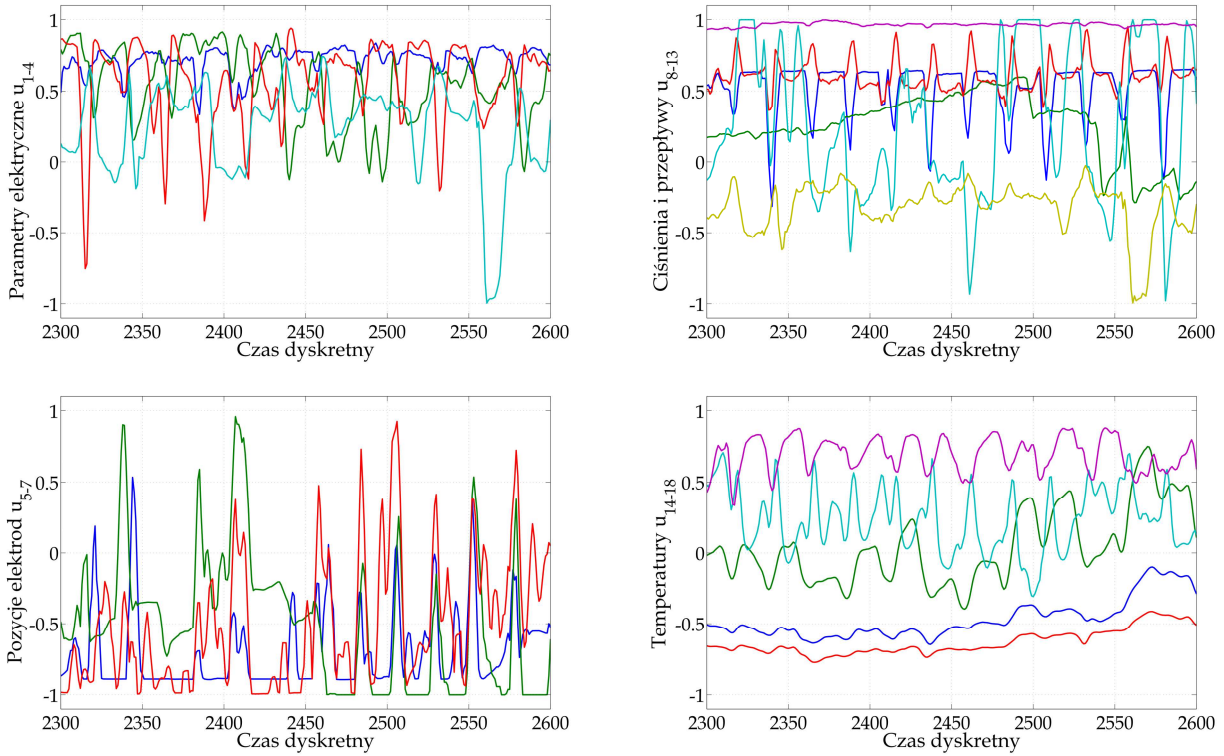
5.2 MODELOWANIE ZŁOŻONEGO PROCESU PRZEMYSŁOWEGO

Kolejny przykład dotyczy zastosowania opracowanego podejścia do modelowania zjawisk zachodzących podczas redukcji miedzi z żużła zawiesinowego. Proces ten realizowany jest przez piec elektryczny do redukcji miedzi z żużła. Piec ten jest istotnym obiektem w ciągu produkcyjnym miedzi [26]. System SCADA podczas monitorowania obiektu rejestruje szereg wielkości fizycznych takich jak: prądy i napięcia, ciśnienia, przepływy, temperatury. Dodatkowo zapisywane są różnego rodzaju informacje dotyczące analizy składu chemicznego żużła zawiesinowego, pozycji elektrod zanurzonych w żużlu, masy wsadu itp. Łącznie gromadzonych jest około 180 zmiennych procesowych. Jednym z ważniejszych parametrów diagnostycznych, które podlegają monitorowaniu, jest temperatura.

Tab. 3. Zestawienie wstępnie wyselekcjonowanych zmiennych procesowych

Zm.	Jed.	Opis zmiennej procesowej
y	°C	Temperatura sklepienia pieca pomiędzy elektrodami E1 i E2
u_1	A	Prąd elektrody E1
u_2	A	Prąd elektrody E2
u_3	A	Prąd elektrody E3
u_4	A	Prąd wentylatora nadmuchu do komory dopalania
u_5	m	Pozycja elektrody E1
u_6	m	Pozycja elektrody E2
u_7	m	Pozycja elektrody E3
u_8	Pa	Ciśnienie w piecu elektrycznym
u_9	MPa	Ciśnienie wody na odpływie z płyt kontaktowych elektrod E2 i E1
u_{10}	Pa	Ciśnienie w komorze dopalania
u_{11}	Pa	Ciśnienie przed chłodnicami
u_{12}	m ³ /h	Przepływ wody do kolektora nr 2, chłodzenie kesonów E1/E2
u_{13}	m ³ /h	Ilość powietrza przed pierwszą sekcją komory dopalania
u_{14}	°C	Temperatura blachy trzonu elektrody E1
u_{15}	°C	Temperatura blachy trzonu elektrody E2
u_{16}	°C	Temperatura blachy trzonu elektrody E3
u_{17}	°C	Temperatura sklepienia pieca pomiędzy elektrodami E2 i E3
u_{18}	°C	Temperatura sklepienia pieca pomiędzy elektrodami E1 i E3

Zadanie, które postanowiono tu rozwiązać, dotyczyło budowy modelu neuronowego temperatury sklepienia pieca pomiędzy elektrodami E1 i E2. Pierwszym etapem budowy modelu było wybranie zmiennych procesowych mających największy wpływ na zmiany wartości tej temperatury. Wstępnie na podstawie dokumentacji technicznej dotyczącej procesu oraz po zasięgnięciu opinii personelu prowadzącego obiekt wyselekcjonowano pewien podzbiór zmiennych procesowych (tab. 3), mogących mieć istotny wpływ na wartości modelowanej temperatury.



Rys. 8. Znormalizowane przebiegi zmiennych niezależnych

Tab. 4. Wybór relewantnych wejść modelu neuronowego

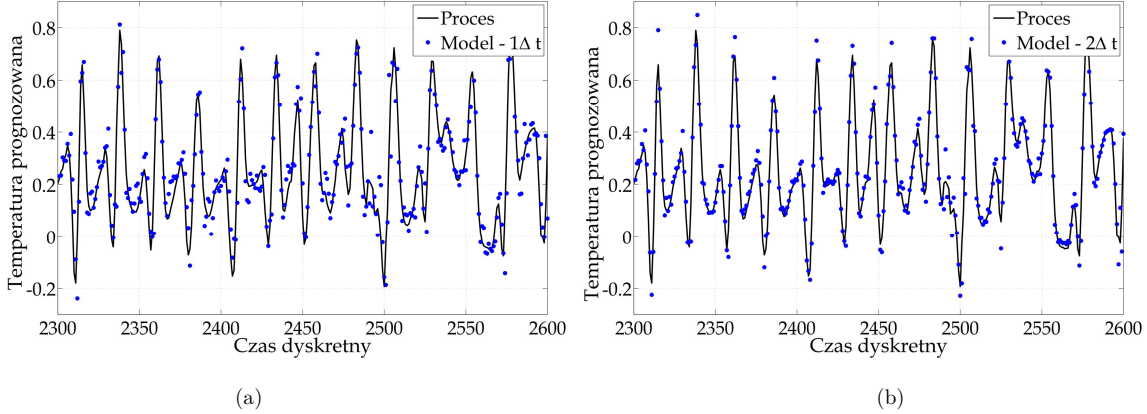
Lp.	Metoda	Parametry metody	Zmienne niezależne									
			u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}	u_{12}	u_{13}
1	MWI	$v_{jk} = I_{jk}, v_{ij} = r_{ij}$	0	<u>1</u>	<u>0.9</u>	0.4	0.4	0	0.1	0	0.4	0
2	MWI	$v_{jk} = nr_{jk}, v_{ij} = r_{ij}$	0	<u>1</u>	<u>0.6</u>	<u>0.6</u>	<u>0.7</u>	0.3	0	0	<u>1</u>	0
3	MWI	$v_{jk} = nr_{jk}, v_{ij} = I_{ij}$	0.4	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	0.4	0.5	<u>1</u>	0.2
4	MWI	$v_{jk} = I_{jk}, v_{ij} = I_{ij}$	0.3	<u>1</u>	<u>1</u>	<u>1</u>	<u>0.7</u>	<u>0.7</u>	0.1	0.1	<u>0.9</u>	0
5	MWI	$v_{jk} = \rho_{jk}, v_{ij} = r_{ij}$	0	0	0.3	0.2	<u>1</u>	<u>1</u>	0.2	<u>0.8</u>	0.4	0
6	MWI	$v_{jk} = \tau_{jk}, v_{ij} = r_{ij}$	0	0	0.3	0.1	<u>1</u>	<u>1</u>	0.2	<u>0.7</u>	0.5	0
7	MWI	$v_{jk} = I_{jk}, v_{ij} = \det_{ij}$	0	<u>1</u>	0.4	0.2	0.2	0.1	0.2	0.1	0	0
8	MWI	$v_{jk} = I_{jk}, v_{ij} = \text{ent}_{ij}$	0.1	<u>0.9</u>	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0
9	MWI	$v_{jk} = I_{jk}, v_{ij} = \text{lam}_{ij}$	0	<u>1</u>	0.5	0.2	0.2	0	0.2	0.2	0	0
10	Γ	$L_2, z = 5$	0.4	0.5	<u>0.6</u>	<u>0.7</u>	<u>1</u>	<u>0.8</u>	<u>0.7</u>	0.5	0.4	<u>0.8</u>
11	Γ	$L_1, z = 5$	<u>0.6</u>	<u>0.6</u>	<u>0.7</u>	<u>0.7</u>	<u>0.6</u>	<u>0.6</u>	<u>0.8</u>	0.4	<u>0.7</u>	<u>0.6</u>
12	Γ	$L_\infty, z = 5$	<u>0.6</u>	<u>0.6</u>	<u>0.7</u>	<u>0.7</u>	<u>0.6</u>	<u>0.6</u>	<u>0.8</u>	0.4	<u>0.7</u>	<u>0.6</u>

* MWI –metoda wskaźników informacyjnych, gdzie I_{jk} - informacja wzajemna, r_{ij} - wsp. korelacji Pearsona, nr_{ij} - nieliniowy wsp. korelacji, ρ_{jk} - wsp. korelacji rang Spearmana, τ_{jk} - wsp. zgodności Kendalla; Γ - test Gamma [9] gdzie L oznacza normę oraz z oznacza liczbę fałszywych sąsiadów użytych podczas analiz.

Tab. 5. Wyniki testowania modeli neuronowych dla różnych konfiguracji zmiennych niezależnych

Model predykcyjny	Struktura	p	Horyzont predykcji			
			$H = 1\Delta t$	$H = 2\Delta t$	$H = 3\Delta t$	
$y(k+H) = f(y, u_{1-3}, u_{5-9}, u_{14-18})$	$14 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	126	3.126	6.059	22.973	x
			0.444	0.473	1.086	
$y(k+H) = f(y, u_{1-3}, u_{5-6}, u_{14-18})$	$11 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	111	3.085	10.944	9.493	x
			0.429	0.558	0.605	
$y(k+H) = f(y, u_{1-3}, u_{5-8}, u_{12}, u_{14-18})$	$14 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	126	2.929	—	25.094	x
			0.442	—	0.844	
$y(k+H) = f(y, u_{1-3}, u_{5-9}, u_{12}, u_{14-18})$	$15 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	131	3.082	6.299	11.251	x
			0.429	0.467	0.667	
$y(k+H) = f(y, u_{1-3}, u_{8-9}, u_{11}, u_{14-18})$	$13 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	116	3.054	6.209	9.304	x
			0.471	0.470	0.611	
$y(k+H) = f(y, u_{1-3}, u_5, u_{14-18})$	$10 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	106	2.992	6.700	9.732	x
			0.421	0.491	0.649	
$y(k+H) = f(y, u_{1-3}, u_{6-10}, u_{13-18})$	$15 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	131	2.824	7.875	24.287	x
			0.429	0.567	0.841	
$y(k+H) = f(y, u_{1-3}, u_{4-10}, u_{12-18})$	$18 \rightarrow 4_{(0,2,0)}^{(2,2,2)} \rightarrow 1$	146	2.917	6.947	23.836	x
			0.452	0.505	0.824	

x - błąd mMAPE; — współczynnik Theila wyznaczony w odniesieniu do prognozy naiwnej



Rys. 9. Prognozowane wartości temperatury sklepienia pieca pomiędzy elektrodami E1 i E2 uzyskane dla modelu neuronowego o najmniejszej liczbie wejść: (a) horyzont $H = 20$ [min.], (b) horyzont $H = 40$ [min.] ($\Delta t = 20$ [min.]

Dane trenujące i testowe utworzono w następujący sposób. Rozpatrywany zbiór danych obejmował około 36 godzin pracy obiektu. Zmienne procesowe rejestrowane były przez system SCADA z krokiem czasowym $\Delta t = 1$ min. Tak zgromadzone dane przetworzono z zastosowaniem filtru cyfrowego o skończonej odpowiedzi impulsowej 40. rzędu. Uzyskane przebiegi poddano kolejno operacji decymacji, wybierając co 20. próbkę sygnału oraz operacji skalowania do przedziału $[-1, 1]$.

W wyniku tych działań otrzymano szeregi czasowe zmiennych procesowych o kroku czasowym $\Delta t = 20$ min. i liczbie próbek równej 2736. Pierwszych 1500 próbek każdego sygnału tworzyło zbiór trenujący. Pozostałe próbki zastosowano na etapie testowania modeli. Przykładowe przebiegi zmiennych niezależnych u_{1-18} pokazano na rys. 8. W kolejnym kroku przyjęto założenie, że modelowana temperatura w głównej mierze zależy od prądów elektrod u_{1-3} oraz temperatur u_{14-18} . Wpływ pozostałych zmiennych procesowych wyznaczono, stosu-

jąc metodę wskaźników informacyjnych (MWI) oraz test Gamma (Γ) [9]. Wyniki obliczeń zapisane są w tab. 4. Biorąc pod uwagę wszystkie wyniki testów zamieszczone w tab. 4, stwierdzono, że najczęściej wskazywanymi zmiennymi, które mają wpływ na temperaturę sklepienia pieca pomiędzy elektrodami E1 i E2 są: pozycje elektrod (u_5, u_6, u_7), ciśnienie w piecu elektrycznym (u_8) oraz ciśnienie wody na odpływie z płyt kontaktowych (u_9).

Dla wszystkich wariantów zmiennych niezależnych utworzono predykcyjne modele neuronowe o stałej strukturze wewnętrznej (dla horyzontów prognozy równych 20, 40 i 60 minut). Podczas dostrajania modeli neuronowych stosowano algorytm hybrydowy według schematu EA-LM.

Przyjęto następujące wartości parametrów funkcji celu (10) dla algorytmu globalnego: $R = 4$, $\lambda = 0.5$, $\omega_0 = 1.5$ oraz dla algorytmu lokalnego: $R = 2$, $\lambda = 0$. W pierwszym etapie treningu stosowano algorytm ewolucyjny (20 epok) z kodowaniem rzeczywisto-

liczbowym chromosomów, z krzyżowaniem heurystycznym ($p_k = 0.8$), z mutacją równomierną z odwzorowaniem Hénona ($r_m = 0.1$), stałą liczebnością populacji równą 100 osobników (populacja bazowa uzyskiwana metodą zaproponowaną w pracy [20]), selekcją proporcjonalną oraz sukcesą elitarną. W kolejnym etapie procesu dostrajania modeli stosowano algorytm Levenberga-Marquardta (10 iteracji) z numerycznie wyznaczanym jakobianem. Uzyskane wyniki modelowania zaprezentowano w tab. 5.

Największą selektywność istotnych zmiennych procesowych uzyskano dla metody wskaźników informacyjnych, dla której współczynnik określający zależność pomiędzy i -tą i j -tą potencjalną zmienną wejściową modelu wyznaczano za pomocą miar bazujących na analizie diagramów rekurencyjnych (17-21). Podobne wyniki otrzymano, gdy jako miary zależności v_{jk} i v_{ij} stosowano odpowiednio informację wzajemną oraz współczynnik korelacji Pearsona. Najmniejszą selektywność uzyskiwano dla metody opartej na teście Γ . Na rys. 9 przedstawiono dokładność uzyskiwanych prognoz dla modelu o najmniejszej liczbie wejść.

6. PODSUMOWANIE

W artykule zaprezentowano sposób modelowania neuronowego procesów technicznych, który znajduje praktyczne zastosowanie w rzeczywistych obiektach przemysłowych.

Literatura

1. Aihara K.: Chaos engineering and its application to parallel distributed processing with chaotic neural networks. In: Proceedings of the IEEE 2002. Vol. 90, Iss. 5.
2. Abarbanel H.: Analysis of observed chaotic data. New York: Springer Verlag, 1996.
3. Awrejcewicz J., Mosdorf R.: Analiza numeryczna wybranych zagadnień dynamik chaotycznej. Warszawa: WNT, 2003.
4. Cun Y. L., Denker J. S., Solla S. A.: Optimal brain damage. „Advances in Neural Information Processing Systems”. Morgan Kaufmann 1990, p. 598 - 605.
5. De Moor B.: DaISy: Database for the identification of systems. Departament of Electrical Engineering, ESAT/SISTA, K.U.Leuven, Belgium, [97-003] Data from a test setup of an industrial winding process. Strona projektu (dn. 05.10.2013): <http://homes.esat.kuleuven.be/~smc/daisy/>.
6. Garfinkel A., Spano M., Ditto W., Weiss J.: Controlling cardiac chaos. „Science” 1992, Vol. 257, p. 1230 - 1235.
7. Hirota K.: Fuzzy-neuro-chaos: research and industrial applications in Japan, „Systems, Man and Cybernetics” 1995, Vol. 3, p. 2446 - 2459.
8. Fradkov A, Evans R.: Control of chaos: methods and applications in engineering. „Annual Reviews in Control” 2005, Vol. 29, p. 33 - 56.
9. Kemp S. E., Wilson I. D., J. Andrew Ware: A tutorial on the gamma test. „International Journal of Simulation Systems, Science and Technology” 2005, Vol. 6 No. 1, p.67 - 75.
10. Korbicz J., Kościelny J. M., Kowalczyk Z., Cholewa, W. (red.): Fault diagnosis: models, artificial intelligence, applications. Berlin, Heidelberg: Springer, 2004.
11. Krishnaiah J., Kumar C., Farugi M.: Modelling and control of chaotic processes through their bifurcation diagrams generated with the help of recurrent neural

Zaproponowano nowy model neuronu będący uogólnieniem jednostek przetwarzających znanych z literatury. Wykazano złożony charakter dynamiki jednostki przetwarzającej charakterystyczny dla tzw. układów chaotycznych. Opisano również metodę uczenia lokalnie rekurencyjnej globalnie jednokierunkowej sieci neuronowej zbudowanej z takich jednostek. Metoda uczenia opracowana została na podstawie tzw. strategii dwufazowej: pierwsza faza uczenia realizowana z zastosowaniem algorytmu ewolucyjnego z mutacją chaotyczną; druga faza realizowana za pomocą metody Levenberga-Marquardta. Opracowano również sposób doboru wejść modelu oraz jego struktury wewnętrznej. Nowum w tym zakresie stanowiło rozwinięcie metody Z. Hellwiga o miary znane z teorii chaosu uwzględniające zależności czasowe w danych.

Praktyczne znaczenie zaproponowanego sposobu modelowania przedstawiono na podstawie danych pozyskanych podczas eksploatacji obiektów rzeczywistych. Uzyskane wyniki potwierdzają duży potencjał opracowanego podejścia.

Badania finansowane częściowo ze środków projektu badawczego promotorskiego nr N N514 3412 33 finansowanego przez MNiSW (kierownik projektu: prof. dr hab. W. Moczulski) oraz częściowo ze środków przeznaczonych na badania statutowe w Instytucie Podstaw Konstrukcji Maszyn Politechniki Śląskiej w Gliwicach.

- network models. Part 2: An industrial study. „Journal of Process Control” 2006, Vol. 16, No. 1, p. 67 - 79.
12. Li Z.: Fuzzy chaotic systems: modeling, control and applications. T. 199 Studies in Fuzziness and Soft Computing. Berlin, Heidelberg: Springer Verlag, 2006.
 13. Li Z., Halang W. A., Chen G.: Integration of fuzzy logic and chaos theory. Studies in Fuzziness and Soft Computing. Berlin, Heidelberg: Springer Verlag, 2006.
 14. Marwan N., Kurths J., Saparin P.: Generalised recurrence plot analysis for spatial data. „Physics Letters A” 2007, Vol. 360, p. 545 - 551.
 15. Marwan N., Romano M. C., Thiel M., Kurths J.: Recurrence plots for the analysis of complex systems. „Physics Reports” 2007, Vol. 438, p. 237 - 32.
 16. Moon F. C.: Chaotic Vibrations. An introduction for applied scientists and engineers. John Wiley and Sons, Inc, 2004.
 17. Moon F. C., Kalmár-Nagy T.: Nonlinear models for complex dynamics in cutting materials. „Phil. Trans. R. Soc. Lond A” 2001, Vol. 359, p. 695 - 711.
 18. Morrison F.: Sztuka modelowania układów dynamicznych - deterministycznych, chaotycznych, stochastycznych. Warszawa: WNT, 1996.
 19. Ott E.: Chaos w układach dynamicznych. Warszawa: WNT, 1997.
 20. Patan K.: Artificial neural networks for the modelling and fault diagnosis of technical processes. T. 377 Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer, 2008.
 21. Przystała P.: Metodyka modelowania neuronowego w diagnostyce procesów z uwzględnieniem elementów teorii chaosu. Gliwice, Pol. Śl., KPKM, 2009. Zesz. nr 141,
 22. Roy R., Murphy T., Maier T., Gills Z., Hunt E.: Dynamical control of a chaotic laser: experimental stabilization of a globally coupled system. „Phys. Rev. Lett.” 1992, Vol. 68, No. 9, p. 1259 - 1262.
 23. Rutkowski L.: Metody i techniki sztucznej inteligencji. Warszawa: Wyd. Nauk. PWN, 2005.
 24. Schuster H.: Chaos deterministyczny: wprowadzenie. Warszawa: Wyd. Nauk. PWN, 1993.
 25. Stavroulakis P.: Chaos applications in telecommunications. CRC Taylor and Francis, 2005.
 26. Szulim R., Moczulski W.: A method of mining knowledge to aid control of complex industrial processes. In: Recent Developments in Artificial Intelligence Methods, W. Cholewa, T. Burczynski i W. Moczulski (eds.). AI-METH Series. Gliwice: Pol. Śl., 2004, p. 273 - 276.
 27. Thiel M., Romano M., Kurths J., Meucci R., Allaria E., Arecchi F.: Influence of observational noise on the recurrence quantification analysis. „Physica D: Nonlinear Phenomena” 2002, Vol. 171, No. 3, p. 138 - 152.
 28. Wiggins S.: Introduction to applied nonlinear dynamical systems and chaos. New York: Springer-Verlag, 2003.
 29. Zadeh L.A.: Fuzzy logic, neural networks and soft computing. „Communications of the ACM” 1994, Vol. 37, No. 3, p. 77 - 84.
 30. Zbilut J., Webber C.: Embeddings and delays as derived from quantification of recurrence plots. „Physics Letters A” 1992, Vol. 171, p. 199 - 203.